



Lopez-Santander, A., & Lawry, J. (2017). An Ordinal Model of Risk Based on Mariner's Judgement. *Journal of the Institute of Navigation*, 70(2), 309-324. <https://doi.org/10.1017/S0373463316000576>

Peer reviewed version

License (if available):
Other

Link to published version (if available):
[10.1017/S0373463316000576](https://doi.org/10.1017/S0373463316000576)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via CUP at <https://www.cambridge.org/core/journals/journal-of-navigation/article/an-ordinal-model-of-risk-based-on-mariners-judgement/251EC346E2B21A8C22E587C800263584>
Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

An Ordinal Model of Risk Based on Mariner's Judgement.

Adan Lopez-Santander¹
Jonathan Lawry²

(Department of Engineering Mathematics, University of Bristol)
(Email: enals@bristol.ac.uk¹, j.lawry@bristol.ac.uk²)

This paper describes a statistical method for learning and estimating the risk posed by other crafts in the vicinity of a vessel and an overview to its possible spatial application, simulating how professional mariners perceive and assess such risk and using navigational data obtained from a standard integrated bridge. We propose a non-linear model for risk estimation which attempts to capture mariners' judgement. Questionnaire data has been collected that captures and quantifies Mariners' judgements of risk for crafts in the vicinity, where each craft is described by measurements that can be obtained easily from the data already present in the ship's navigational equipment. The dataset has then been used for analysis, training and validating Ordered Probit Models in order to obtain a computationally efficient data driven model for estimating the risk probability vector posed by other crafts. Finally, we discuss how this risk model can be incorporated into decision making and path finding algorithms.

KEY WORDS

1. Craft's Risk. 2. Risk estimation. 3. Mariner's knowledge. 4. Collision avoidance

1. INTRODUCTION. There is an extensive existing body of research in the field of collision avoidance as an aid for navigators in manned crafts and also for use in autonomous unmanned crafts, e.g. (Belkhouche and Bendjilali, 2013, Lambert et al., 2008, Plamen Angelov and Michael Everett, 2008). Amongst this research, there is a broad consensus concerning the needs to acquire a precise representation of the environment surrounding the craft and, most importantly, of processing the acquired data for assessing risk of collision before any decision can be made. Some deterministic approaches compute the risk of collision as the rate of change of the relative bearing to crafts in the vicinity, either indirectly, by considering the Distance at Closest Point of Approach (DCPA) and Time to Closest Point of Approach (TCPA) as in the appraisal index proposed by Kearon (Kearon, 1977), or by directly computing the derivative of the relative bearing to the craft, e.g. Angelov and Everett (Plamen Angelov and Michael Everett, 2008).

This is also the approach recommended by the International Regulations for Prevention of Collisions at Sea, COLREG (International Maritime Organization (IMO), 1972), which in rule 7, part D, defines risk of collision as a function of time and bearing as follows:

“In determining if risk of collision exists the following considerations shall be among those taken into account:

- (i) such risk shall be deemed to exist if the compass bearing of an approaching vessel does not appreciably change;
- (ii) such risk may sometimes exist even when an appreciable bearing change is evident, particularly when approaching a very large vessel or a tow or when approaching a vessel at close range.”

Current industry standards use a DCPA-TCPA concrete relationship alarm system in line with the above where arbitrary minimum thresholds are set for DCPA and TCPA. Should a target’s DCPA and TCPA trespass such limits, an alarm will be given (IMO, 2004).

Hilger and Baldauf (Hilgert and Baldauf, 1997) proposed a rule based concrete model to standardise the meaning of risk given by the COLREG using data provided by Automatic Radar Plotting Aid (ARPA) and defining four crisp risk classes. Alternatively, numerous proposals to add a layer of fuzzy logic to the risk model and or actions to avoid collisions have been suggested. For example, Bukhari et al (Bukhari et al., 2013) propose a fuzzy model to capture the relationship between DCPA, TCPA and change in bearing in order to assist VTS centres making decisions. In addition, Perera et al (Perera et al., 2012) have applied both fuzzy inference and Bayesian methods in order to assess risk of collision and to take evasive action and Goerland et al (Goerlandt et al., 2015) present a comprehensive rule based expert system with fuzzy inference where the knowledge domain has been defined by consultation with professional mariners. Furthermore, a number of probabilistic approaches have been proposed that take account of unknown factors in order to predict risk and possible trajectories; see (Belkhouche and Bendjilali, 2013) or (Lambert et al., 2008) and Simsir et al (Simsir et al., 2014) for the application of Artificial Neural Networks for predicting positions of vessels in a collision alert system. These methods add a layer of sophistication to the geometric approach. Chin and Debnath (Chin and Debnath, 2009) describe an ordinal model of risk based on data from a survey of pilots, again taking account of the relationship between DCPA and TCPA and incorporating day or night navigation and ship tonnage as an indicator of manoeuvrability.

Common to all these studies is the direct association of risk with the spatial possibility of a collision or the conceptualisation of risk as a consequence of collision only. For these methods, a craft that is navigating in parallel to the observing vessel would not account for risk, as for instance in a traffic separation scheme as defined by the International Maritime Organisation (IMO, 2013). However, the navigator’s judgement of the risk posed by a neighbouring craft seems to involve other factors than just the geometric calculation of trajectories (Goerlandt and Montewka, 2015, Curtis, 1986) . Furthermore, in many to many scenarios, where many crafts are at risk of colliding with one another, perception of risk seems to become even more complex and depends on how different crafts interact with each another and on their particular idiosyncrasies. Hence, crafts can be perceived as posing a risk for the mariner even when a collision is not imminent or even impossible.

In this paper, we present a machine learning approach to assessing inherent risk for crafts in the vicinity of a vessel. This approach will infer a model deriving from human judgement and experience, which provides a distribution of risk levels for a

given context that can be incorporated as a cost function in navigation or avoidance algorithms. This work is a part of a larger research project in collision avoidance in which it is a key part of the intelligent risk assessment in navigation and collision avoidance algorithms.

To elicit information on the mariners' judgement of risk, a questionnaire has been designed to record data on a number of scenarios. The questionnaire was completed by 425 professional marine navigators, who were asked to assess the risk of individual crafts in different scenarios with one or many crafts and used an ordinal ranking scale from 1 to 5 to quantify their perceived level of risk. A non-linear regression model, Ordered Probit, has then been used to learn and to estimate risk for crafts in new scenarios based on a number of attributes. The resulting estimation has the form of a probability vector for a distribution of risk over the values 1 to 5. The objective is not to find an exact value for risk of collision but to learn an operational model of the inherent risk of a neighbouring craft in a given scenario based on the experience of the professional mariner.

An outline of the remainder of the paper is as follows: Section 2 describes in detail the method chosen to elicit the data necessary to define and train our model. The model is then presented and explained in section 3. Section 4 discusses the performance of the model in comparison with a Naïve-Bayes classifier and proposes a possible application of the resulting estimation. Finally, the conclusions, section 5, discuss briefly the limitations of our model and suggest directions for further research.

2. QUESTIONNAIRE DESIGN AND DATA COLLECTION. A professional mariner's judgement of risk develops through training and through experience gained in many different encounters with other crafts during navigation. This implicit knowledge of good seamanship can only be acquired by training, practice and actual experience. With the objective of capturing the mariner's judgement of neighbouring crafts' risk, a questionnaire was designed so as to provide a craft's risk ranking on a scale from 1, very low, to 5, very high, for different scenarios. The questionnaire was composed of two parts: The first part presented single craft situations, one-on-one encounters between the mariner's vessel and a neighbour craft, and described them textually with the following characteristics being provided: Closest Point of Approach (CPA), Time to Closest Point of Approach (TCPA), Relative Situation, Colour, and Trajectory Variability (see Figure 1). For this first part, no graphical representation was presented.

The set of variables was selected through consultation with professional mariners. DCPA, presented to the mariner as CPA, and TCPA are spatiotemporal relational values between a given craft and the vessel being navigated by the mariner and are functions of bearing, range, speed and course of the two crafts involved; DCPA being the closest distance between two objects should the trajectories not change and TCPA the time to reach such point. Relative Situation can take any of the values defined by the COLREG (International Maritime Organization (IMO), 1972): Head on, crossing, overtaking and or overtaken. Colour can be either green, red or undefined when it is not clear or cannot be determined, and is the method used by mariners to relatively position themselves with respect to another craft's side; either on the starboard or port side of the reference craft respectively. The Trajectory

Variability of a craft indicates the confidence in the prediction of its future positions, thus, affecting the certainty in DCPA and TCPA. The professional mariner does this intuitively by observing how erratically a craft's trajectory is. The questionnaire simplifies this to a Boolean variable i.e. erratic or not erratic. In a later application of this research, Kalman filters are used to process the positional data of crafts to predict trajectories, and a threshold on the resulted error covariance matrix of the state estimate is used to determine the value of this Boolean variable in real time. Particularities of the individual crafts were not presented in this questionnaire, hence, variables like tonnage or manoeuvrability, that have been considered in other works, are not contemplated.

The set of variables are deemed to be inherent to the craft and not dependent on the environment. Accordingly, the questionnaire did not include geographical features, visibility constraints –it was assumed to occur in good visibility- or weather conditions, for example, as these were presumed to be common to all nearby crafts in a scenario and are hence assumed constants. How these variables would affect the absolute risk is beyond the scope of this paper. The COLREG convention is a fundamental pillar of maritime training and as such is always part of any assessment made in collision avoidance at sea. Ideally, we would have liked our data to have been independent of this set of rules, for our wish is to learn inherent risk of crafts which we could then use to apply any set of rules, and we have made this clear when the questionnaire was presented to the mariners. But our results suggest that COLREG may still have influenced the participants' judgement of risk, for example in their different assessments of a green or red craft in most cases. The model presented in this paper aims to learn about inherent risk of other crafts that could then be used as a variable to account when applying the COLREG at a higher layer.

This first part of the questionnaire consisted of 100 different scenarios from which 20 were randomly selected and presented to each participant. This approach allowed us to collect data on a wider range of scenarios while at the same time limiting the length of the sessions.

Question 2 out of 20

You are in a Crossing situation with a ship which has a CPA of 1.5 nm and a TCPA of 2 min. Her aspect is Red and her course seems to be steady.

Situation: Crossing Aspect: Red CPA: 1.5 nm TCPA: 2 min Course: steady

Answer 2 out of 20

Which value of risk, from 1 to 5, would you assign to this target?

Please, select a value from the list - 1 very low, 5 very high- to assign a risk to the target

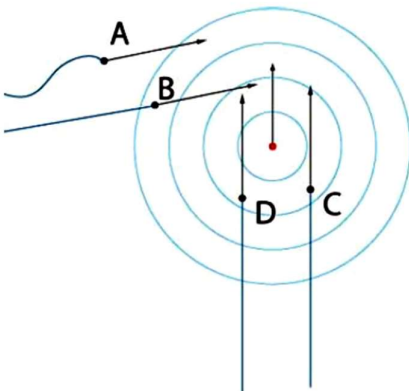
Situation Risk: 1
2
3
4
5

Previous Next

Figure 1. Presentation of textual questions.

The second part of the questionnaire was comprised of simple graphical representations of scenarios, including single craft, one-on-one, and multi-craft, many-on-many situations. The questions were accompanied by a concise supporting text with the objective of eliminating possible ambiguity resulting from the simplicity of the graphics. The graphical scenarios were depicted as a simple polar coordinate representation of a Radio Detection And Ranging (RADAR) or Electronic Chart Display and Information System (ECDIS) screens, showing neighbouring crafts with vectors for speed and course and a tracked trajectory. It was not explicitly specified if the course and speed vectors were true or relative (see Figure 2), however, the responses show a unimodal distribution with low variance suggesting that a common interpretation has been adopted across the participants. We note that a number of the scenarios presented in this part of the questionnaire corresponded to the graphical representation of some of the text based questions included in the first section.

Question 2 out of 10
 In the following scenario, where your ship is represented by the red dot; which values of risk would you assign to the different targets? also, which value of risk would you give to the whole scenario?



Your instruments are giving you the following information for each target.

- A: Speed 9kn; CPA 1; TCPA 6 min
- B: Speed 9kn; CPA 0.5; TCPA 4 min
- C: Speed 9kn; CPA 1; TCPA 6 min
- D: Speed 9kn; CPA 0.7; TCPA 7 min

Answer 2 out of 10
 Please, select a value from the lists -1 very low, 5 very high- to assign a risk to the targets and to the scenario

Target A Risk: Target B Risk: Target C Risk: Target D Risk: Scenario Risk:

[Previous](#) [Next](#)

Figure 2. Presentation of graphical questions.

In total there were 50 different graphical scenarios from which a subset of 10 were randomly selected to present to each participant for assessment. They were then asked to evaluate the risk for each craft represented and the overall risk for the scenario itself. Other works present a textual representation of risk in line with the International Maritime Organisation recommendations for alerts in Integrated Navigation Systems (IMO, 2007); see Hilgert and Baldauf (Hilgert and Baldauf, 1997), Goerlandt et al. (Goerlandt et al., 2015) and Chin and Debnath (Chin and Debnath, 2009) for examples. However, in our work the participants had to respond to the questions using a numeric scale ranging from 1 to 5 rating the risk posed by each craft. This scale and range gave adequate resolution and at the same time helped the participants to identify an appropriate risk level with equally spaced categories. Also, it avoided any semantic confusion associated with the use of labels and its interpretation; see Wildt and Mazis for a discussion on the latter, Schwarz and Clark for a study on how interactions between labels and rating scales modify the meaning and Preston and Colman for an study in optimal number of categories. (Wildt and Mazis, 1978, Norbert Schwarz and Clark, 1991, Pohl, 1981, Preston and Colman, 2000).

The questionnaire was distributed online and it was accessible anonymously through the University of Bristol website. The server's software stored the responses directly into a database and logged them with a random session identification number. The participants were recruited mainly through maritime organizations, shipping companies and training centres providing over 8000 observations over the period in which the questionnaire was active. The sample size for each question/scenario varies but on average is 39, with a minimum of 24 and a maximum of 65 participants.

3. DATA ANALYSIS. As expected, there is variation between the participants regarding their judgement of the risk posed by the crafts in a given scenario. This may be due to differences in experience or understanding or may simply be natural variation between individuals. After a first analysis of the data, we observed an underlying normal distribution on the responses to the questions and since neither of the continuous independent variables used satisfies the properties of superposition for the dependent variable, risk, then this suggests that there is a non-linear relationship between them and a normally distributed error. This, added to the ordered scale proposed for the responses, suggested the use of an Ordered Probit.

3.1. Use of Ordered Probit Model. The Ordered Probit Model has its roots in bio-statistics (Aitchison and Silvey, 1957) and was introduced into the social sciences by MacKelvey and Zavoina (McKelvey and Zavoina, 1975). It has often been used since for the analysis and prediction of dependent ordered variables with an underlying, continuous and nonlinear metric. Indeed, there are many applications of the Ordered Probit model, ranging from accident injury prediction (O'Donnell and Connor, 1996) to estimation of customer's enjoyment of films by the Netflix Prize's winner (Koren, 2009, Andreas Toscher, 2009, Piotte and Chabbert, 2009) and has previously been used by Chin and Debnath to model risk of collision in port water navigation (Chin and Debnath, 2009). The model assumes an underlying linear relationship characterized by

$$Y = X\beta + \varepsilon \quad (1)$$

where Y is the continuous latent unobserved variable, X is a vector of independent variables defining the data, β is an unknown coefficient for x and ε is a random disturbance term which is assumed to be normally distributed according to $\sim N(0,1)$. The central idea is that a latent real variable is underlying the ordinal set of responses and that we observe these responses instead of the latent variable. The real line is then divided into variable regions that represent the ordinal categories; in our case five regions. The observed ordinal variable Z takes the values $1, \dots, 5$ corresponding to the set of risk categories $z \in \{1, \dots, 5\}$.

$$Z = z \Leftrightarrow \mu_{z-1} < Y \leq \mu_z \quad (2)$$

where μ are the bounds of the regions on Y which define the values of Z as intervals of the real line. The probability of observing a particular ordinal outcome in an Ordered Probit model is therefore given by:

$$P[Z = z] = \Phi(\mu_z - X\beta) - \Phi(\mu_{z-1} - X\beta) \quad (3)$$

where Φ is the cumulative normal distribution and $z = 1, \dots, 4$ and, for $z = 5$, as ε is assumed to be multivariate normal with $\sigma^2 = 1$, then:

$$P[Z = 5] = 1 - \Phi(\mu_4 - X\beta) \quad (4)$$

The β parameters and the μ bounds of the model are estimated by means of Maximum Likelihood Estimation (MLE), normally using Newton-Raphson method on the log-likelihood function, which for the Ordered Probit model is:

$$\log(L) = \sum_{i=1}^N \log[\Phi(\mu_z - x_i\beta) - \Phi(\mu_{z-1} - x_i\beta)] \quad (5)$$

where x is the value of vector X in the sample i , for a sample of size N . For further details regarding this method see McKelvey and Zavoina or William E. Becker and Peter E. Kennedy (McKelvey and Zavoina, 1975, Becker and Kennedy, 1992).

3.2. Independent Variables for estimating risk. Let Z be the perceived risk, which has the format of an ordinal polychotomous variable taking on values from 1 to 5. For each craft, there will be a probability that its perceived risk takes each of the values from 1 to 5. Let R_i be the vector of probabilities of perceived risk for the given i 'th craft defined by the vector of independent variables X_i :

$$R_i = \langle P(Z = 1 | X_i), P(Z = 2 | X_i), \dots, P(Z = 5 | X_i) \rangle \quad (6)$$

The selection of independent variables for the model is intended to isolate the craft from the environment and only those descriptive of the craft's relations with others

crafts in the neighbourhood are considered. Also, particulars of the craft such as size, type of vessel or navigational circumstances, such as method of propulsion, are not contemplated for this research. Each craft is assumed to be a two dimensional point in the Euclidean plane. The initial set of independent variables was chosen by consultation with professional mariners as described in table 1.

Table 1. Initial set of independent variables.

Variable	Description
Continuous variables	
DCPA	Closest Point of Approach
TCPA	Time to Closest Point of Approach
DCPA TCPA	Interaction between DCPA and TCPA
Speed	True speed of the craft in knots
Relative Speed	Relative speed of the craft in knots
Relative Course	Relative course of the craft in degrees
Range	Range to the craft
Bearing	Bearing to the craft
Aspect	The aspect of the craft in degrees
Estimated position error	The error of the craft's Kalman filters estimation. Used for real time application.
Discrete variables (dichotomous)	
Red	1 if the aspect of craft is red; 0 otherwise
Green	1 if the aspect of craft is green; 0 otherwise
Course Erratic	1 if the craft's track is erratic; 0 otherwise
Course Variable	1 if the craft's track is not steady; 0 otherwise
Altering course	1 if there is a broad alteration of course; 0 otherwise
Head On	1 if the encounter situation is 'head on' as defined by the COLREG; 0 otherwise
Crossing	1 if the encounter situation is 'crossing' as defined by the COLREG; 0 otherwise
Overtaking	1 if the encounter situation is 'overtaking' as defined by the COLREG; 0 otherwise
Trajectory Variability	1 if the craft's trajectory is erratic; 0 otherwise

We have compared a number of different models using diverse selections of variables from the list in table 1 in terms of Schwarz Bayesian Information Criterion (SBIC) (Schwarz, 1978) and by calculating Average Marginal Effects of each variable for each value from the possible outcome scale. As a result of this process we selected a final model which has the continuous variables DCPA and TCPA and the discrete (dichotomous) variables Red, Green, Course Erratic, Head on, Crossing, Overtaking. This model offers a balance between goodness of fit and economy of calculation for real time applications.

The results in table 2 show that DCPA has large effect on the perceived risk, which concurs with the established methods for assessing risk of collision. However, TCPA on its own seems to have little impact on the risk assessments of the navigators in our study. This will be discussed in more detail in the next section of this paper

3.3. Parameter estimates. Maximum likelihood estimates of the structural parameters β and the bounds μ on Y are shown in table 2. Average Marginal

effects of the independent variables, the discrete change in probability for each of the values of a given i th variable averaged across the observed values of the rest of the variables in the model (Bartus, 2005), are shown in table 3 for all possible outcomes of our risk scale. Note that Average Marginal Effects in non-linear models are not constant and should only be taken in consideration as an indicator and not an estimator (Ai and Norton, 2003, Greene, 2010).

Table 2. Parameter estimation for the independent variables and μ bounds on Y

VARIABLES	Parameter Estimate β
Continuous Variables	
DCPA	-0.395***, (0.00855)
TCPA	-0.0249***, (0.00166)
Discrete variables (dichotomous)	
Red	0.555***, (0.128)
Green	0.445***, (0.132)
Course Erratic	0.522***, (0.0253)
Head On	0.269***, (0.0449)
Crossing	-0.337**, (0.132)
Overtaking	-0.0654, (0.0437)
μ Bounds	
-	
μ_1	-1.273***, (0.0361)
μ_2	-0.622***, (0.0348)
μ_3	-0.0292, (0.0342)
μ_4	0.493***, (0.0345)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3. Average Marginal Effects of variables for different outcomes

VARIABLES	AVERAGE MARGINAL EFFECTS				
	Risk = 1	Risk = 2	Risk = 3	Risk = 4	Risk = 5
Continuous Variables					
DCPA	0.0878*** (0.00163)	0.0338*** (0.00109)	0.00684*** (0.000836)	-0.0199*** (0.000827)	-0.109*** (0.00234)
TCPA	0.00553*** (0.000368)	0.00213*** (0.000146)	0.000431*** (5.84e-05)	-0.00125*** (9.45e-05)	-0.00684*** (0.000454)
Dummy variables (dichotomous)					
RedColor	-0.123*** (0.0284)	-0.0475*** (0.0109)	-0.00961*** (0.00240)	0.0279*** (0.00655)	0.152*** (0.0349)
GreenColor	-0.0988*** (0.0294)	-0.0381*** (0.0113)	-0.00770*** (0.00240)	0.0224*** (0.00674)	0.122*** (0.0362)
ErraticCourse	-0.116***	-0.0446***	-0.00903***	0.0263***	0.143***

	(0.00566)	(0.00228)	(0.00108)	(0.00161)	(0.00672)
HeadOn	-0.0598***	-0.0230***	-0.00466***	0.0135***	0.0739***
	(0.00999)	(0.00385)	(0.000934)	(0.00231)	(0.0123)
Crossing	0.0750**	0.0289**	0.00584**	-0.0170**	-0.0927**
	(0.0293)	(0.0113)	(0.00234)	(0.00670)	(0.0362)
Overtaking	0.0145	0.00560	0.00113	-0.00329	-0.0180
	(0.00969)	(0.00375)	(0.000773)	(0.00220)	(0.0120)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Some of our previous assumptions are supported by the estimation of the β parameters but there are also surprises, as in the case of the low absolute value of the parameter for variable TCPA.

As expected, DCPA has a significant effect in the perceived risk; the large negative parameter estimate indicates a significant increase of perceived risk as the value of DCPA decreases, which is consistent with the message that a value 0 for a craft's DCPA anticipates a collision. As mentioned, this is also the approach of well established methods for assessing risk of collision i.e. a DCPA with value 0 is equivalent to a steady bearing with a craft. However, our data and the TCPA's parameter estimate gives us a different insight. One might intuitively think that a smaller TCPA would also have a significant effect, increasing significantly the perceived risk, but this is not clear from its estimated parameter value. The low absolute value of the latter suggest that only a slight increase in perceived risk results from a decrease in TCPA.

Another parameter that has a large effect on the perceived risk is the steadiness or uncertainty of the trajectory of the craft, included in our model as a discrete dichotomous variable. For an erratic course, the perceived risk increases notably as indicated by the high estimated parameter value.

Furthermore, table 2 gives a relatively high absolute value to the HeadOn variable, suggesting that head on encounters are judged to have more risk than crossing or overtaking encounters. In this case, it is possible that the COLREG is affecting the response and this situation creates more uncertainty in the mariner than the others, for there is not a 'stand on' or 'give way' vessel defined in the rules or a clear Colour of the crafts. This contrasts with the other two situations where it is made clear in the COLREG which vessel must 'give way' in relation to their Colours; the 'head on' situation requires an avoiding action from the two crafts involved. Red aspect crafts, those on the starboard bow of the craft observing, are also perceived as carrying more risk. Again, this may be the effect of the COLREG on the mariners' judgement as the rules will always give them 'stand on' rights in normal visibility situations. This seems to leave the burden for action on the mariner and a judgement of risk perhaps associated with this responsibility.

4. RESULTS. The performance of the model has been evaluated using ten-fold cross validation (Kohavi, 1995) over the original questionnaire's responses data set and

has been compared against the well-known Gaussian Naïve-Bayes classifier (Hand and Yu, 2001).

4.1. *Benchmarking with Naïve-Bayes.* Naïve-Bayes classifiers are a popular class of probabilistic classifiers that make use of Bayes' theorem with strong assumptions of independence between the variables used, see Hand and Yu for an interesting description of its effectiveness (Hand and Yu, 2001) or Rish for an empirical perspective (Rish, 2001). It is important to note that we are using Gaussian Naïve-Bayes as a probability estimator and not as a classifier. There is evidence in the literature (Lowd and Domingos, 2005) to suggest that Naïve-Bayes is in general a reasonably accurate and efficient general approach and hence provides a good benchmark for our proposed model.

We use the Kullback-Liebler divergence (Kullback and Leibler, 1951) between observed and predicted distributions as a measure to compare the performance of the models and also a simple Euclidean distance between the predicted data during cross-validation and observed data. The Kullback-Leibler divergence, or relative entropy, is a non-symmetrical measure of the difference between two probability distributions expressed by:

$$D_{KL}(P \parallel Q) = \sum_i \ln \left(\frac{P(i)}{Q(i)} \right) P(i) \quad (7)$$

This metric is widely used in Information Theory since its results can be interpreted as the average number of extra units of information required to encode data generated by one distribution, P , using coding from a different distribution, Q .

Figure 3 shows the results of this test with the values D_{KL} in the Y axis and the scenario in the X axis. It can be clearly seen that the Ordered Probit model significantly outperforms the Gaussian Naïve-Bayes model on predicting risk for new crafts using our dataset. The values D_{KL} , in Nats, for every scenario depicted in the questionnaire are better (lower values better) in almost every case and in providing consistent and homogeneous results.

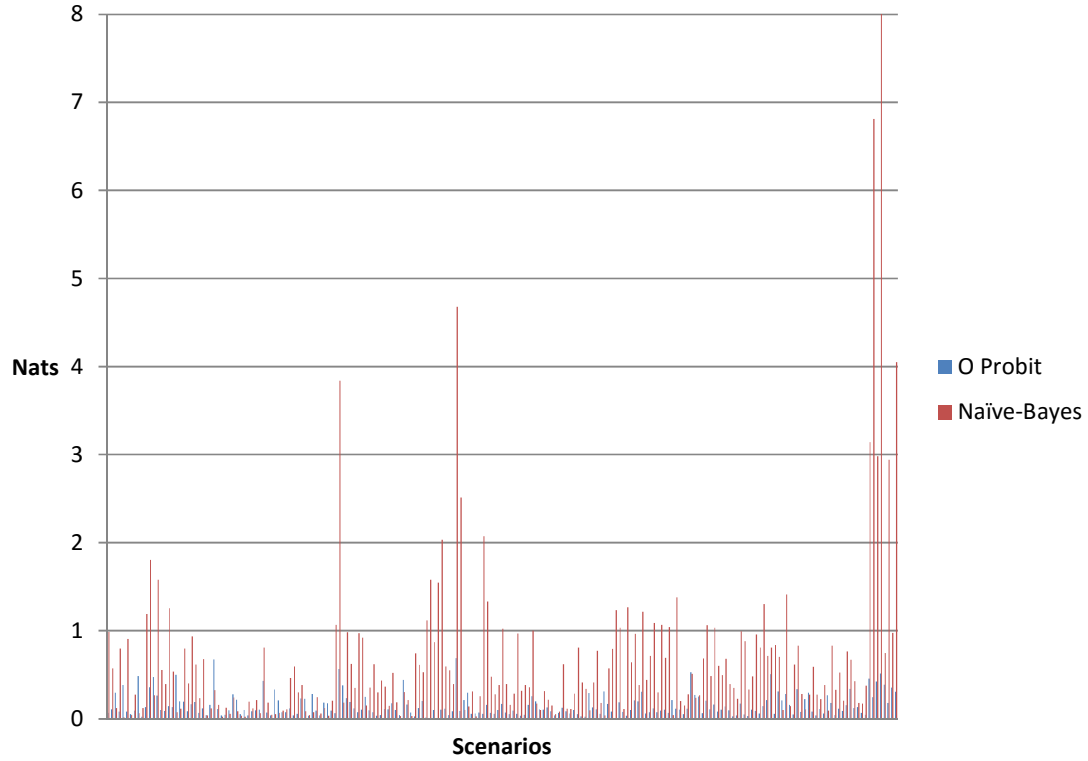


Figure 3. Nats for each scenario for O Probit and Naïve-Bayes.

Figure 4 shows the Euclidean distance between the observed data and the Ordered Probit model predictions in the cross validation and figure 5 that for the Gaussian Naïve-Bayes model, where it is clear the larger dispersion of the predictions for the latter. The line $x = y$ would represent a perfect fit of a model.

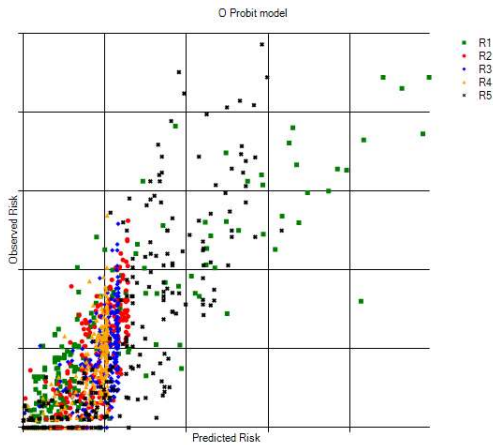


Figure 4.

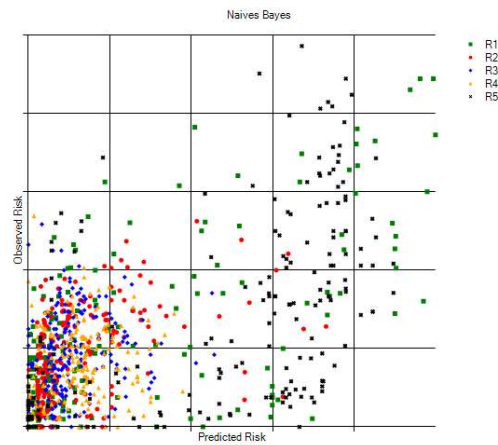


Figure 5.

4.2. *Apply the Risk Model.* Using the presented model as a risk estimator provides a vector of probabilities representing the inherent risk distribution of any given craft. We propose a method to embed such vector of risk in the navigation space where applicable. The resulting estimation can be converted into a spatiotemporal risk cost function by means of nested areas representing different

risk levels for a given craft. This cost function can thereupon be employed to define risk shaped pseudo-static ‘obstacles’ incorporated into all sorts of path finding algorithms. We present a three dimensional example of this application where the radius of such areas is set to constant and their height is defined by the predicted probability for the level of risk. Note that it could also be possible to work with two dimensional areas of variable radius defined by the risk:

Let the craft Tg have a circular domain of an arbitrary diameter at time zero t_0 . With a simple projection of the craft’s domain to a possible point of collision given Tg’s vector and our own craft’s (myObject) speed, we can calculate a likely area in a time interval during which our craft could potentially be invading Tg’s domain if a course between V_1 and V_2 were taken (see Figure 6). Course V_c would invariably lead to a collision at a future time should the course and speed of both, craft Tg and myObject, not change.

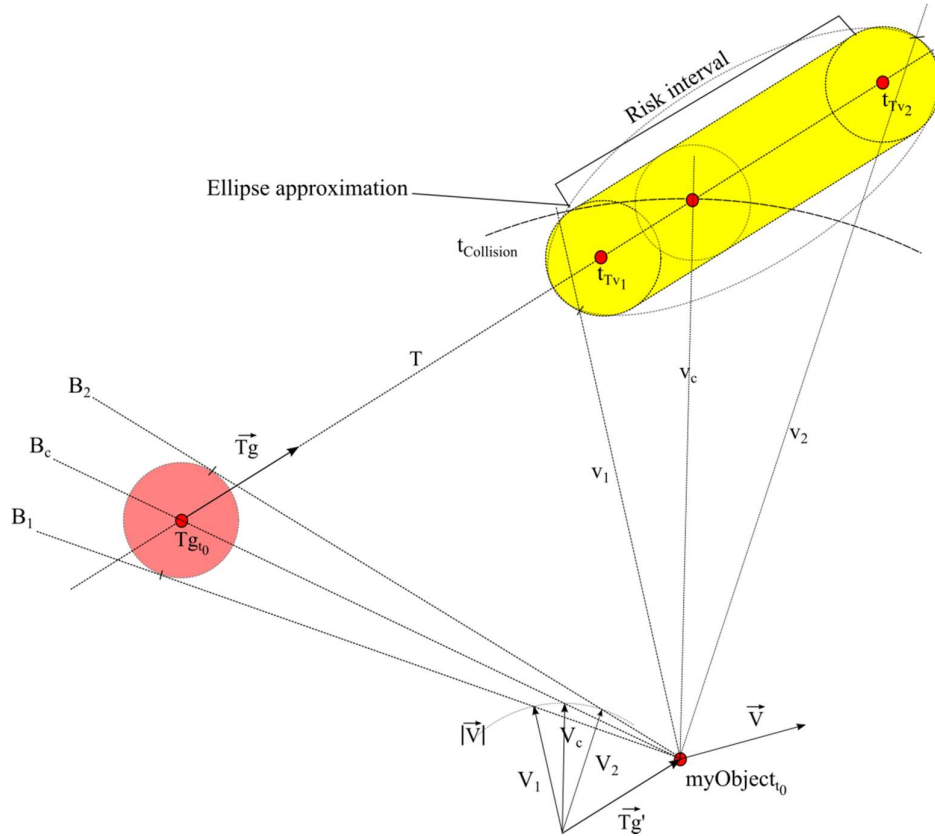


Figure 6. Projection of a given craft’s domain to a collision time.

At t_0 , the craft’s estimated risk is mapped concentrically at equal distances, being risk 1 at the centre and risk 5 at the periphery. The risk acquired when crossing a risk interval for a given craft can be easily calculated in the actual related craft’s domain. Thus, the trajectory that receives the higher risk is the one that would collide with the target, which crosses the five zones at its maximum chord or its diameter.

managed in order to successfully navigate through them. We show a simple example of application for mapping the obtained risk vector to the space-time and a cost function to be used in path finding and optimisation algorithms.

Humanlike understanding of a craft's risk in its complexity, including quantifying uncertainty, offers a powerful tool for Intelligent Navigation Systems. The analysis in this paper is obviously limited to the training dataset collected and further development, including a continuous learning capability, would be necessary for real world applications. This first approach offers a rather simple model which should be expanded to include possible interactions between neighbouring crafts, their explicit changes in speed or course, i.e. somehow contemplated in trajectory variability, size of craft and rate of turn i.e. implying manoeuvrability, for instance.

The dataset obtained in our questionnaire can potentially offer responses to some of these variables, interactions for instance, but does not contain enough information to be able to learn from crafts' size or manoeuvrability. Eliciting new data with a new questionnaire would be desirable to further advance in learning and modelling risk.

REFERENCES

- AI, C. & NORTON, E. C. 2003. Interaction terms in logit and probit models. *Economics Letters*, 80, 123-129.
- AITCHISON, J. & SILVEY, S. D. 1957. The Generalization of Probit Analysis to the Case of Multiple Responses. *Biometrika*, 44, 131-140.
- ANDREAS TOSCHER, M. J. 2009. The BigChaos Solution to the Netflix Grand Prize. AT&T Labs.
- BARTUS, T. 2005. Estimation of marginal effects using margeff. *Stata Journal*, 5, 309-329.
- BECKER, W. E. & KENNEDY, P. E. 1992. A Graphical Exposition of the Ordered Probit. *Econometric Theory*, 8, 127-131.
- BELKHOUCHE, F. & BENDJILALI, B. 2013. Dynamic collision risk modeling under uncertainty. *Robotica*, 31, 525-537.
- BUKHARI, A. C., TUSSEYEVA, I., LEE, B.-G. & KIM, Y.-G. 2013. An intelligent real-time multi-vessel collision risk assessment system from VTS view point based on fuzzy inference system. *Expert Systems with Applications*, 40, 1220-1230.
- CHIN, H. C. & DEBNATH, A. K. 2009. Modeling perceived collision risk in port water navigation. *Safety Science*, 47, 1410-1416.
- CURTIS, R. G. 1986. A Ship Collision Model for Overtaking. *The Journal of the Operational Research Society*, 37, 397-406.
- GOERLANDT, F. & MONTEWKA, J. 2015. Maritime transportation risk analysis: Review and analysis in light of some foundational issues. *Reliability Engineering & System Safety*, 138, 115-134.
- GOERLANDT, F., MONTEWKA, J., KUZMIN, V. & KUJALA, P. 2015. A risk-informed ship collision alert system: Framework and application. *Safety Science*, 77, 182-204.
- GREENE, W. 2010. Testing hypotheses about interaction terms in nonlinear models. *Economics Letters*, 107, 291-296.
- HAND, D. J. & YU, K. 2001. Idiot's Bayes: Not So Stupid after All? *International Statistical Review / Revue Internationale de Statistique*, 69, 385-398.
- HILGERT, H. & BALDAUF, M. 1997. A common risk model for the assessment of encounter situations on board ships. *Deutsche Hydrografische Zeitschrift*, 49, 531-542.
- IMO 2004. SOLAS V, Annex 34. Resolution MSC.192(79).

IMO 2007. Adoption of the Revised Performance Standards for Integrated Navigation Systems (Ins). In: COMMITTEE, T. M. S. (ed.). IMO.

IMO 2013. *Ships' Routeing*, London, IMO Publishing.

INTERNATIONAL MARITIME ORGANIZATION (IMO) 1972. International Regulations for Prevention of Collisions at Sea. London: IMO.

KEARON, J. Computer program for collision avoidance and track keeping. Conference on Mathematics Aspects of Marine Traffic, 1977. 229-242.

KOHAVI, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*. Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc.

KOREN, Y. 2009. The BellKor Solution to the Netflix Grand Prize.

KULLBACK, S. & LEIBLER, R. A. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22, 79-86.

LAMBERT, A., GRUYER, D. & PIERRE, G. S. A fast Monte Carlo algorithm for collision probability estimation. Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on, 17-20 Dec. 2008 2008. 406-411.

LOWD, D. & DOMINGOS, P. Naive bayes models for probability estimation. Proceedings of the Twentysecond International Conference on Machine Learning, 2005. 529-536.

MCKELVEY, R. D. & ZAVOINA, W. 1975. A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology*, 4, 103-120.

NORBERT SCHWARZ, B. K., HANS-J. HIPPLER, ELISABETH NOELLE-NEUMANN, & CLARK, L. 1991. Rating Scales: Numeric Values May Change the Meaning of Scale Labels. *The Public Opinion Quarterly*, 55, 570-582.

O'DONNELL, C. J. & CONNOR, D. H. 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. *Accident Analysis & Prevention*, 28, 739-753.

PERERA, L. P., CARVALHO, J. P. & SOARES, C. G. 2012. Intelligent Ocean Navigation and Fuzzy-Bayesian Decision/Action Formulation. *IEEE Journal of Oceanic Engineering*, 37, 204-219.

PIOTTE, M. & CHABBERT, M. 2009. The Pragmatic Theory solution to the Netflix Grand Prize. Pragmatic Theory Inc.

PLAMEN ANGELOV, C. D. B., COSTAS XIDEAS, CHARLES PATCHETT, DAREN ANSELL, & MICHAEL EVERETT, G. L. 2008. A Passive Approach to Autonomous Collision Detection and Avoidance in Uninhabited Aerial Systems. *Tenth International Conference on Computer Modeling and Simulation*.

POHL, N. F. 1981. Scale Considerations in Using Vague Quantifiers. *The Journal of Experimental Education*, 49, 235-240.

PRESTON, C. C. & COLMAN, A. M. 2000. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychol (Amst)*, 104, 1-15.

RISH, I. An empirical study of the naive Bayes classifier. IJCAI-01 workshop on "Empirical Methods in AI", 2001.

SCHWARZ, G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6, 461-464.

SIMSIR, U., AMASYALI, M. F., BAL, M., ÇELEBI, U. B. & ERTUGRUL, S. 2014. Decision support system for collision avoidance of vessels. *Applied Soft Computing*, 25, 369-378.

WILDT, A. R. & MAZIS, M. B. 1978. Determinants of Scale Response: Label versus Position. *Journal of Marketing Research*, 15, 261-267.